

ANALISIS SENTIMEN MAHASISWA TERHADAP FASILITAS UNIVERSITAS TELKOM MENGGUNAKAN METODE JARINGAN SARAF TIRUAN DAN TF-IDF

Muhammad Farhan Muzakki¹, Jondri², Rian Febrian Umbara³

^{1,2,3}Fakultas Informatika, Universitas Telkom, Bandung

¹zakkifarhan@student.telkomuniversity.ac.id, ²jondri@telkomuniversity.ac.id,

³rianum@telkomuniversity.ac.id

Abstrak

Keberadaan kuisioner dapat membantu instansi manapun untuk meningkatkan fasilitas ataupun kinerjanya. Tetapi dengan kuisioner yang berjumlah ratusan bahkan ribuan akan menyulitkan instansi untuk mengetahui kesimpulan dari seluruh data kuisioner. Data kuisioner yang diambil sebagai acuan yaitu data yang berisi tanggapan positif dan negatif. Oleh karena itu, penelitian ini menganalisis kuisioner mengenai fasilitas di Universitas Telkom. Analisis dilakukan dengan melakukan klasifikasi kuisioner yang berisi sentimen mahasiswa tentang fasilitas di Universitas Telkom. Metode klasifikasi yang digunakan dalam penelitian ini adalah Jaringan Saraf Tiruan (JST) dengan model *Multi Layer Perceptron* yang dikombinasikan dengan fitur ekstraksi untuk dapat mendeteksi negasi dan pembobotan menggunakan *Term Frequency – Inverse Document Frequency (TF-IDF)*. Hasil pengujian pada aplikasi yang dibangun memperlihatkan bahwa akurasi memberikan tingkat akurasi yang baik hingga 91,23%.

Kata kunci: analisis sentimen, jaringan saraf tiruan, multi layer perceptron, TF-IDF

Abstract

The existence of a questionnaire can help any agency to improve its facilities or performance. But with hundreds or thousands of questionnaires it will be difficult for agencies to find conclusions from all questionnaire data. Questionnaire data taken as a reference are data that contain positive, negative, and neutral responses. Therefore, this study tried to analyze the questionnaire regarding facilities at Telkom University. The analysis was carried out by classifying questionnaires containing student sentiments about facilities at Telkom University. The classification method used in this study is Artificial Neural Networks (ANN) with a Multi Layer Perceptron model combined with extraction features to be able to detect negation and weighting using *Term Frequency – Inverse Document Frequency (TF-IDF)*. The test results on the applications that are built show that accuracy provides a good level of accuracy, and can be implemented.

Keywords: sentiment analysis, artificial neural network, multi layer perceptron, TF-IDF

1. Pendahuluan

Latar Belakang

Universitas Telkom menggunakan kuisioner sebagai umpan balik dari mahasiswa untuk mengetahui keadaan fasilitas kampus. Kuisioner wajib di jawab oleh mahasiswa sebagai salah satu persyaratan mencetak kartu ujian sebelum masa cetak kartu ujian berakhir. Sekarang data kuisioner diolah secara manual oleh Admin. Admin berusaha menganalisis jawaban kuisioner untuk menentukan fasilitas apa saja yang dianggap mahasiswa sebagai fasilitas yang bernilai positif dan fasilitas mana saja yang dianggap mahasiswa bernilai negatif. Sehingga dapat menjadi pertimbangan dari pihak kampus untuk melakukan perawatan dan memperbaiki kerusakan.

Dengan adanya kuisioner tersebut maka fasilitas di kampus dapat terpantau secara keseluruhan dari umpan balik mahasiswa. Tetapi Admin masih melakukan analisa secara manual. Sehingga perawatan dan perbaikan tidak dapat dilaksanakan secara tepat waktu. Dengan adanya permasalahan di atas maka dapat diatasi dengan dibangun sistem yang dapat menentukan kuisioner positif, negatif, dan netral. Serta objek apa saja yang mendapatkan umpan balik positif, dan negatif. Untuk membangun sistem yang dapat menentukan sentimen tersebut maka dibutuhkan metode yang dapat membantu menganalisis hasil kuisioner tersenut.

Untuk melakukan klasifikasi sentimen dari opini berbahasa Indonesia yang ada, diperlukan metode klasifikasi yang mampu mengklasifikasikan opini menjadi 2 kelas klasifikasi yaitu sentimen positif dan sentimen negatif. Pada tugas akhir ini digunakan metode klasifikasi Jaringan Saraf Tiruan (JST) untuk melakukan pengklasifikasian opini berbahasa Indonesia. Jaringan Saraf Tiruan (JST) merupakan salah satu metode klasifikasi yang meniru cara kerja otak manusia yang dapat menyelesaikan suatu permasalahan dengan metode pembelajaran

(*learning*). Kelebihan JST salah satunya adalah kemampuannya dalam beradaptasi sehingga mampu belajar dari data masukan yang diberikan sehingga dapat memetakan hubungan antara masukan dan keluarannya [1].

Topik dan Batasannya

Berdasarkan latar belakang yang sudah dipaparkan, penelitian ini akan membahas mengenai penentuan kelas opini positif dan negatif dari hasil kuisioner Universitas Telkom dengan menggunakan metode *Multi Layer Perceptron* dengan algoritma *Backpropagation* dan untuk mengetahui parameter yang dapat mempengaruhi nilai akurasi dari metode *Multi Layer Perceptron* dengan algoritma *Backpropagation*. Terdapat rumusan masalah pada penelitian ini yaitu untuk mengetahui pengaruh dari jumlah *Hidden layer*, *Learning rate*, *Dropout*, dan pengacakan data terhadap akurasi yang dihasilkan dalam menentukan opini positif dan negatif terhadap fasilitas yang disediakan oleh Universitas Telkom menggunakan metode *Multi Layer Perceptron*. Kemudian pada penelitian ini terdapat batasan masalah, batasan masalah yang pertama pada penelitian ini adalah *dataset* terdiri dari 10.000 data kuisioner yang didapatkan dari mahasiswa Universitas Telkom. Batasan masalah yang kedua adalah data kuisioner dihasilkan dari kuisioner tanggal 11 November 2016 mengenai seluruh fasilitas mencakup layanan akademik, sistem informasi dan asrama.

Tujuan

Tujuan dari penelitian ini adalah untuk mengetahui parameter yang dapat menentukan tingkat akurasi dari komputasi berdasarkan metode Jaringan Saraf Tiruan (JST) dan *Term Frequency – Inverse Document Frequency (TF-IDF)* dan mengetahui hasil data rekaman kuisioner mahasiswa terhadap fasilitas di Universitas Telkom setelah diolah menggunakan metode Jaringan Saraf Tiruan (JST).

Organisasi Tulisan

Berikutnya pada bagian 2 membahas studi terkait pada penelitian ini, bagian 3 membahas perancangan sistem yang dilaksanakan pada penelitian ini, bagian 4 membahas evaluasi dan bagian 5 membahas kesimpulan pada penelitian ini.

2. Studi Terkait

2.1 Analisis Sentimen

Analisis sentimen mempelajari pandangan individu, evaluasi, tingkah laku dan perasaan terhadap orang, individu, masalah, aktivitas dan subjek [2]. Sentimen analisis yakni proses pembelajaran, memahami, mengekstran dan mengolah data yang berupa tekstual untuk mendapatkan informasi [3]. Tujuan dari sentimen analisis adalah memutuskan sikap atau pendapat dari seseorang terhadap suatu topik atau target. Analisis sentimen dapat digunakan untuk menentukan nilai kesukaan atau ketidaksukaan seseorang terhadap suatu barang. Nilai yang sering digunakan pada sentimen analisis yaitu positif, dan negatif. Nilai ini dapat digunakan untuk dijadikan parameter dalam pengambilan keputusan.

2.2 Text Mining

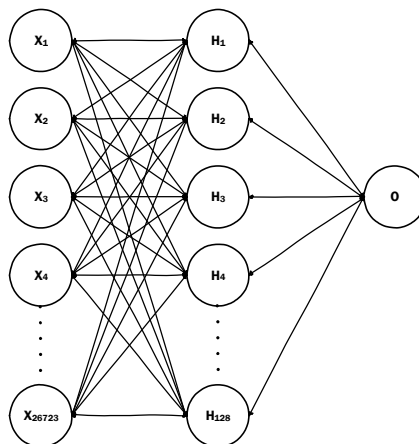
Text mining bertujuan untuk menemukan pola yang tersembunyi pada sumber tertentu agar dapat digunakan untuk suatu tujuan [4]. Untuk melakukan analisis pembuat keputusan, maka dibutuhkan data tidak terstruktur yang berjumlah besar dalam bentuk dokumen. *Text mining* bukanlah sebuah fungsi, akan tetapi kumpulan dari berbagai macam fungsi yang dikombinasikan dan disebut fungsi *Text mining*. Fungsi utama dari *Text mining* meliputi *Searching*, *Information Extraction*, *Categorization*, *Summarization*, *Prioritization*, *Clustering*, *Information Monitor*, *Question & answers* [5].

Text mining merupakan bagian dari *Data mining*, akan tetapi tahapan proses pada *Text mining* lebih banyak dibanding tahapan proses pada *Data mining*, karena data teks memiliki data yang tidak terstruktur sehingga perlu dilakukan beberapa tahap yang pada intinya mengubah data menjadi lebih terstruktur. Berdasarkan ketidakaturan struktur data teks, maka proses *text mining* memerlukan beberapa tahap awal yang pada intinya adalah mempersiapkan agar teks dapat diubah menjadi lebih terstruktur. Tahapan *text mining* yang dilakukan secara umum adalah tahap *Case Folding*, *Remove Punctuation*, *Tokenization*, *Stopword Removal*, dan *Lemmatization*. [6]

2.3 Term frequency – inverse document frequency

Term frequency – inverse document frequency atau biasa sering disebut *TF-IDF* adalah metode pembobotan kata dengan menghitung nilai *TF* dan juga menghitung kemunculan sebuah kata pada koleksi dokumen teks secara keseluruhan [6]. Metode ini menggabungkan 2 konsep perhitungan bobot yaitu frekuensi kemunculan sebuah kata di dalam sebuah dokumen tertentu dan *inverse* frekuensi dokumen yang mengandung kata tersebut [7]. *Inverse document frequency (IDF)* adalah jumlah dokumen yang mengandung sebuah *term* didasarkan pada seluruh dokumen yang ada pada data set.

2.4 Jaringan Saraf Tiruan



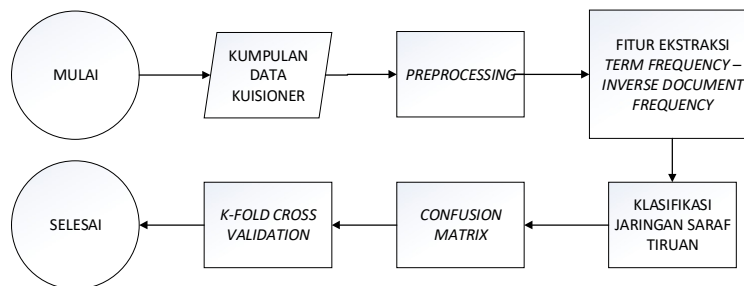
Gambar 1. Visualisasi Multi Layer Perceptron

Berdasarkan Gambar 1, Jaringan Saraf Tiruan (JST) adalah suatu jaringan yang memodelkan sistem saraf otak manusia (*neuron*) dalam melaksanakan tugas pengenalan pola, khususnya klasifikasi [1]. JST merupakan suatu sistem yang didasarkan pada cara kerja jaringan saraf manusia. JST memungkinkan suatu sistem belajar dan melakukan generalisasi sehingga diharapkan sistem tidak hanya mengenali data-data yang sudah pernah diberikan, tetapi juga data baru [8].

3. Sistem yang Dibangun

3.1 Diagram Blok Sistem

Pada tahapan ini, akan dijelaskan proses yang terjadi pada sistem. Dapat dilihat pada diagram blok pada Gambar 2:



Gambar 2. Visualisasi Gambaran Umum Sistem

3.2 Dataset

Dataset adalah sebuah himpunan yang didalamnya berisi data-data. Data yang digunakan adalah data kuisioner fasilitas mahasiswa Universitas Telkom. Penelitian ini menggunakan 10.000 hasil kuisioner yang berisi opini positif dan negatif.

3.3 Preprocessing

Preprocessing yaitu tahap dimana data dari dokumen teks yang tidak terstruktur diolah menjadi data yang terstruktur. *Preprocessing* merupakan proses awal pada teks untuk mempersiapkan teks agar dapat menjadi data yang dapat diolah lebih lanjut [9]. Berikut adalah penjelasan tahapan preprocessing:

1. Case Folding

Pada tahap ini, mengubah penggunaan huruf kapital menjadi bentuk standar atau huruf kecil atau *lowercase*. Contohnya sebagai berikut :

Sebelum *Case Folding*: Fasilitas kampus sudah bagus dan baik!

Sesudah *Case Folding*: fasilitas kampus sudah bagus dan baik!

2. Remove Punctuation

Pada tahap ini, menghilangkan tanda baca dalam teks. Contohnya sebagai berikut :

Sebelum *Remove Punctuation*: fasilitas kampus sudah bagus dan baik!

Sesudah *Remove Punctuation*: fasilitas kampus sudah bagus dan baik

3. Tokenization

Tokenization merupakan tahap yang dilakukan setelah *Case folding*, dimana pada tahap ini merupakan proses pemotongan *String input* berdasarkan tiap kata yang menyusunnya. *Tokenization* juga akan memecah sekumpulan karakter dalam suatu teks ke dalam suatu kata. Contohnya sebagai berikut :

Sebelum *Tokenization*: fasilitas kampus sudah bagus dan baik

Sesudah *Tokenization*: fasilitas, kampus, sudah, bagus, dan, baik

4. Stopword Removal

Proses penghilangan kata yang sering muncul pada dokumen, karena tidak memberikan ciri yang signifikan untuk membedakan dokumen satu dengan yang lain. Contohnya adalah “dan” atau “sebuah”. Karena pada faktanya kata yang mencirikan suatu dokumen biasanya frekuensi kemunculannya jarang. Sehingga untuk mempersingkat proses klasifikasi, kita dapat menghilangkan *Stopword* tersebut. Contohnya sebagai berikut:

Sebelum *Stopword Removal*: fasilitas, kampus, sudah, bagus, dan, baik

Sesudah *Stopword Removal*: fasilitas, kampus, bagus, baik

5. Lemmatization

Lemmatization adalah proses untuk menemukan bentuk dasar dari sebuah kata. Contohnya sebagai berikut:

Sebelum *Lemmatization*: membanggakan

Sesudah *Lemmatization*: bangga

3.4 Term frequency – inverse document frequency

Pada tahap ini, sistem akan menghitung jumlah kemunculan *term*, *inverse document frequency* (idf) dan juga *term weighting* pada dokumen *dataset* menggunakan proses *TF-IDF*.

$$idf = \log \left(\frac{n}{df_i} \right) \quad (1)$$

Kemudian untuk proses pembobotan dari *term* yang ada menggunakan rumus sebagai berikut :

$$w_{t,d} = tf_{t,d} \cdot \log \frac{n}{df_i} \quad (2)$$

3.5 Multi Layer Perceptron

Multi Layer Perceptron (MLP) adalah arsitektur jaringan yang memiliki banyak lapisan. MLP memiliki satu atau lebih *hidden layer* [1]. Salah satu algoritma pelatihan MLP adalah *Backpropagation* (Propagasi Balik). Cara kerja Algoritma *Backpropagation* adalah mengoreksi kesalahan dan memperbaikinya [10]. Algoritma *Backpropagation* dilakukan dalam 3 tahap yaitu propagasi maju, propagasi mundur, dan perubahan bobot. Berikut adalah algoritma *Backpropagation* [1]:

1. Mendefinisikan matriks *Input*, *Hidden*, *Output*.
2. Menginisialisasikan arsitektur jaringan, *Learning rate*, dan nilai bobot melalui nilai acak dengan interval nilai sembarang dengan interval [0, 1].
3. Pelatihan Jaringan
 - Propagasi Maju
Dengan bobot yang telah diacak pada tahap ke dua. Menghitung keluaran dari *Hidden layer* berdasarkan persamaan berikut (menggunakan fungsi aktivasi *Relu*) [11]:

$$Z_{net_j} = v_{jo} + \sum_{i=1}^n x_i v_{ji} \quad (3)$$

$$z_j = \left(z_{net_j} \right) = \max (0, z_{net_j}) \quad (4)$$

Hasil keluaran *Hidden layer* (z_j) digunakan untuk mendapatkan nilai *Output layer* dengan persamaan berikut (menggunakan fungsi aktivasi *Sigmoid*) [11]:

$$y_{net_k} = w_{ko} + \sum_{j=1}^p z_j w_{kj} \quad (5)$$

$$y_k = f(y_{net_k}) = \frac{1}{1 + e^{-y_{net_k}}} \quad (6)$$

- Propagasi Mundur

Melakukan perhitungan factor δ unit *Output* berdasarkan *Error* pada setiap unit *Output* dengan persamaan berikut[11]:

$$\delta_k = (t_k - y_k) f'(y_{net_k}) = (t_k - y_k) y_k (1 - y_k) \quad (7)$$

Selanjutnya melakukan perhitungan suku perubahan bobot w_{kj} dengan laju percepatan a dengan persamaan berikut[11]:

$$\Delta w_{kj} = a \cdot \delta_k \cdot z_j \quad (8)$$

Kemudian melakukan perhitungan δ unit tersembunyi berdasar pada error pada setiap unit tersembunyi dengan persamaan berikut[11]:

$$\delta_{net_j} = \sum_{k=1}^m \delta_k w_{kj} \quad (9)$$

Faktor δ unit tersembunyi dapat dipresentasikan dengan persamaan berikut[11]:

$$\delta_j = \delta_{net_j} f'(Z_{net_j}) = \delta_{net_j} z_j \quad (10)$$

Kemudian melakukan perhitungan suku perubahan bobot v_{ji} dengan persamaan berikut[11]:

$$\Delta v_{ji} = a \cdot \delta_j \cdot x_i \quad (11)$$

- Perubahan Bobot

Melakukan perhitungan seluruh perubahan bobot yang menuju unit output dengan persamaan berikut[11]:

$$w_{kj}(\text{baru}) = w_{kj}(\text{lama}) + \Delta w_{kj} \quad (11)$$

Melakukan perhitungan seluruh perubahan bobot yang menuju unit tersembunyi dengan persamaan berikut[11]:

$$v_{ji}(\text{baru}) = v_{ji}(\text{lama}) + \Delta v_{ji} \quad (12)$$

4. Tahapan diatas adalah untuk satu kali siklus pelatihan. Pelatihan harus diulang-ulang hingga jumlah siklus tertentu.

3.6 Validasi

3.6.1 K-Fold Cross Validation

K-Fold Cross Validation digunakan untuk memvalidasi model yang telah dibangun. Jenis *Cross Validation* yang digunakan pada penelitian ini adalah *5-Fold Cross Validation* dengan cara membagi *Data testing* dan *Data training* dengan perbandingan 2:8. Dengan menggunakan *5-Fold Cross Validation* diharapkan dapat menghasilkan nilai akurasi terbaik.

3.6.2 Confusion Matrix

Tabel 1. Visualisasi *Confusion Matrix*

#	PREDIKSI	
REALITA	TP	FN
	FP	TN

Berdasarkan Tabel 1, *Confusion Matrix* dipilih sebagai metode validasi karena data yang diperoleh memiliki jumlah opini positif dan negatif yang tidak seimbang.

Terdapat 4 istilah pada *Confusion Matrix* yaitu:

1. TP adalah *True Positive*, yaitu jumlah data positif yang terklasifikasi dengan benar oleh sistem.
2. TN adalah *True Negative*, yaitu jumlah data negatif yang terklasifikasi dengan benar oleh sistem.
3. FN adalah *False Negative*, yaitu jumlah data negatif namun terklasifikasi salah oleh sistem.
4. FP adalah *False Positive*, yaitu jumlah data positif namun terklasifikasi salah oleh sistem

$$Akurasi = \frac{TP+TN}{TP+TN+FP+F} * 100\% \quad (13)$$

$$Presisi = \frac{TP}{FP+TP} * 100\% \quad (14)$$

$$Recall = \frac{TP}{FN+TP} * 100\% \quad (15)$$

Berdasarkan nilai *True Negative (TN)*, *False Positive (FP)*, *False Negative (FN)*, dan *True Positive (TP)* dapat diperoleh nilai akurasi, presisi dan *recall*. Nilai akurasi menggambarkan seberapa akurat sistem dapat mengklasifikasikan data secara benar. Dengan kata lain, nilai akurasi merupakan perbandingan antara data yang terklasifikasi benar dengan keseluruhan data. Nilai presisi menggambarkan jumlah data kategori positif yang diklasifikasikan secara benar dibagi dengan total data yang diklasifikasi positif. Sementara itu, *recall* menunjukkan berapa persen data kategori positif yang terklasifikasikan dengan benar oleh sistem. [12]

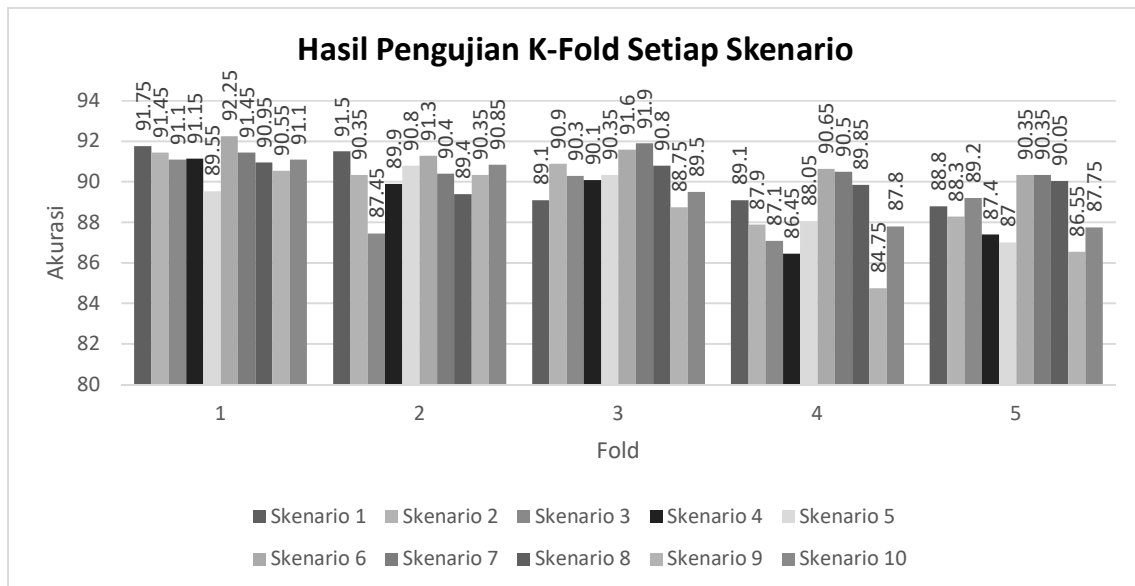
4. Evaluasi

4.1 Hasil Pengujian

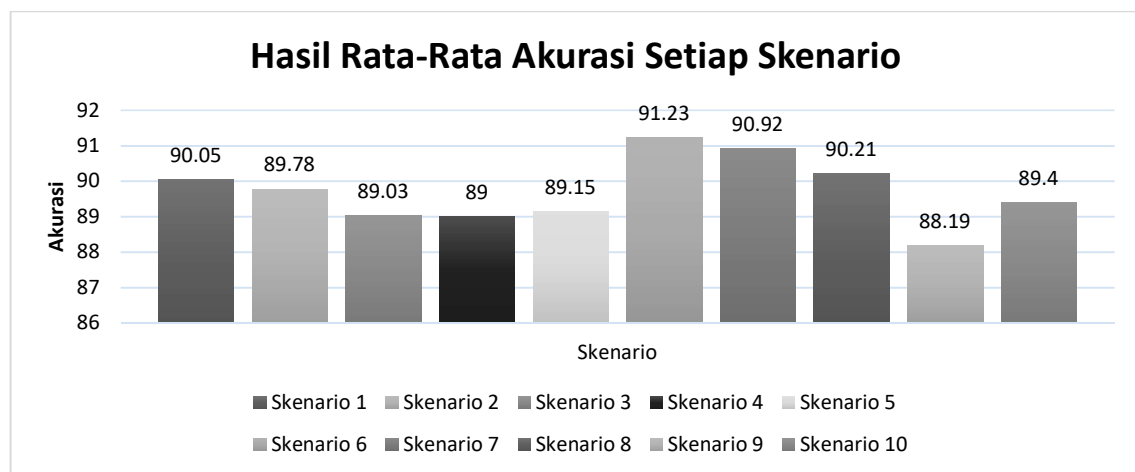
Pengujian menggunakan metode *K-Fold Cross Validation* dengan *K* yang berjumlah 5. Setiap *Fold* terdapat 10 skenario. Setiap skenario memiliki jumlah *Learning Rate*, *Hidden Layer*, *Drop Out*, *Shuffle* yang berbeda-beda. Tabel berikut adalah hasil rata-rata dari seluruh fold pada setiap model:

Tabel 2. Hasil Rata-Rata Pengujian

Model	Skenario					Hasil Rata-Rata				
	Learning Rate	Data Shuffling	Hidden Layer	Drop Out	Output Layer	Train Accuracy	Test Accuracy	Recall	Precision	Accuracy
ANN_01	0.001	X	64	X	1	96.05	90.05	96.21	91.01	90.05
ANN_02	0.001	X	64	X	1	95.79	89.78	96.72	90.30	89.78
ANN_03	0.001	X	64	0.2	1	93.92	89.03	97.43	88.95	89.03
ANN_04	0.001	O	64	X	1	95.04	89	97.10	89.16	89
ANN_05	0.001	O	64	0.2	1	94.49	89.15	97.11	89.30	89.15
ANN_06	0.005	O	64	X	1	97.92	91.23	93.95	94.27	91.23
ANN_07	0.01	O	64	X	1	98.39	90.92	92.98	94.76	90.92
ANN_08	0.1	O	64	X	1	98.37	90.21	92.20	94.54	90.21
ANN_09	0.0005	O	64	X	1	93.92	88.19	97.30	88.20	88.19
ANN_10	0.001	O	128	X	1	95.58	89.4	96.72	89.89	89.4



Gambar 3. Grafik Hasil Pengujian Pada Setiap Fold Disetiap Skenario



Gambar 4. Grafik Hasil Rata-Rata Akurasi Setiap Skenario

4.2 Analisis Hasil Pengujian

Dari tabel 2 dapat disimpulkan bahwa skenario ke 7 (tujuh) dengan Learning rate 0.01, menggunakan pengacakan data, menggunakan Hidden Layer sebanyak 64 Layer mendapatkan nilai akurasi terbaik pada perhitungan Train Accuracy dengan persentase 98.39% dan nilai Precision terbaik dengan persentase 94.76%.

Skenario ke 6 (enam) dengan Learning rate 0.005, menggunakan pengacakan data, menggunakan 64 Hidden layer mendapatkan nilai akurasi terbaik pada perhitungan Test Accuracy dengan persentase 91.23% dan nilai Accuracy menggunakan perhitungan Confusion Matrix terbaik dengan persentase 91.23%.

Lalu pada Skenario ke 3 (tiga) dengan Learning rate 0.001, tidak menggunakan pengacakan data, menggunakan Drop Out sebanyak 20% mendapatkan nilai recall terbaik dengan persentase 97.43%.

Dari Gambar 3 dapat disimpulkan bahwa terjadinya fluktuasi dari akurasi yang didapatkan oleh setiap skenario disetiap Fold. Setiap Fold memiliki Data Train dengan opini positif dan negative yang berbeda-beda. Maka dari itu terjadi fluktuasi terhadap akurasi. Sejak Fold pertama hingga Fold terakhir, Skenario ke 6 (enam) mendapatkan hasil akurasi yang stabil.

Dari seluruh hasil pengujian dapat diketahui bahwa jumlah *Hidden layer*, *Learning rate*, *Drop out*, dan pengacakan data berpengaruh terhadap nilai akurasi. Besarnya *Hidden layer* tidak menjamin akurasi menjadi lebih baik dibandingkan dengan jumlah *Hidden layer* yang lebih sedikit. Perbandingan opini negatif dan positif pada *dataset* sangat berpengaruh terhadap hasil pengelompokan opini positif dan negatif.

5. Kesimpulan

Berdasarkan seluruh hasil pengujian dan analisis yang telah dilakukan pada penelitian tugas akhir ini, maka dapat ditarik kesimpulan dan saran sebagai berikut:

Kesimpulan

1. Jumlah Node pada Hidden Layer, Input Layer, nilai *Learning Rate*, penggunaan *Drop Out*, *Data Shuffling* berpengaruh pada nilai akurasi yang dihasilkan, Skenario 6 dengan Learning rate 0.005, menggunakan pengacakan data, menggunakan 64 Hidden layer mendapatkan nilai akurasi terbaik pada perhitungan Test Accuracy dengan persentase 91.23% dan nilai Accuracy menggunakan perhitungan Confusion Matrix terbaik dengan persentase 91.23%.
2. Jumlah Node pada Hidden Layer berpengaruh terhadap akurasi dapat dilihat dari hasil Skenario 4 dan 10. Hidden Layer yang lebih besar akan meningkatkan nilai akurasi walaupun tidak signifikan, Skenario 10 mendapatkan nilai akurasi 89.5% sedangkan Skenario 4 mendapatkan nilai akurasi 69%.
3. Semakin tinggi nilai *Learning Rate* semakin besar nilai akurasi dapat dilihat dari hasil Skenario 4 dan 6. Skenario 6 mendapatkan nilai akurasi 91.23% sedangkan Skenario 4 mendapatkan nilai akurasi 89%.
4. Penggunaan *Drop Out* juga berpengaruh terhadap akurasi dapat dilihat di Skenario 4 mendapatkan nilai akurasi 89% dan Skenario 5 yang menggunakan *Drop Out* mendapatkan nilai akurasi 89.15%.

Saran

1. Penelitian ini dapat dikembangkan untuk kebutuhan instansi untuk menganalisis hasil dari umpan balik mahasiswa.
2. Proses menganalisis data umpan balik memakan waktu yang sangat lama, sehingga akan lebih baik apabila menggunakan *platform* yang lain sebagai *compiler*. (Contoh: *MatLab*)
3. Penelitian ini dapat dikembangkan apabila dihubungkan dengan *Website* instansi sehingga mahasiswa dapat melihat hasil dari analisis umpan balik.

Daftar Pustaka

- [1] M. S. Dr. Suyanto, S.T., *Data Mining*. Informatika Bandung, 2017.
- [2] Basari *et al.*, "Opinion Mining of Movie Review using Hybrid Method of Support Vector Machine and Particle Swarm Optimization," *Procedia Eng.* 53, pp. 453 – 462, 2013.
- [3] D. Bandorski *et al.*, "Contraindications for video capsule endoscopy," *World J. Gastroenterol.*, vol. 22, no. 45, pp. 9898–9908, 2016.
- [4] Indriati and Ridok, "Sentiment analysis For Review Mobile Applications Using Neighbor Method Weighted K-Nearest Neighbor (NWKNN)," *J. Environ. Eng. Sustain. Technol.*, 2016.
- [5] Mustafa, Akbar, and Sultas, *Knowledge Discovery using Text Mining: A Programmable Implementation on Information Extraction and Categorization*. Pakistan, 2009.
- [6] A. Achmad and A. A. Ilham, "Implementasi Algoritma Term Frequency – Inverse Document Frequency dan Vector Space Model untuk Klasifikasi Dokumen Naskah Dinas," vol. 257, pp. 88–92, 2012.
- [7] A. A. Maarif, "Penerapan Algoritma TF-IDF untuk Pencarian Karya Ilmiah," *Dok. Karya Ilm. | Tugas Akhir | Progr. Stud. Tek. Inform. - SI | Fak. Ilmu Komput. | Univ. Dian Nuswantoro Semarang*, no. 5, p. 4, 2015.
- [8] D. F. Paman, M. A. Muslim, and M. Dhofir, "Analisis Implementasi Jaringan Syaraf Adaptif Untuk Peramalan Kebutuhan Energi Listrik Wilayah Malang," *Neutrino*, vol. 2, no. 2, pp. 117–133, 2010.
- [9] Falahah and D. D. A. Nur, "DATA PREPROCESSING UNTUK ANALISIS SENTIMEN BERBASIS LEKSIKON UNTUK PREDIKSI CALON KEPALA DAERAH JAWA BARAT TAHUN 2018 PADA MEDIA SOSIAL TWITTER," 2015.
- [10] Yusran, "IMPLEMENTASI JARINGAN SYARAF TIRUAN (JST) UNTUK MEMPREDIKSI HASIL NILAI UN MENGGUNAKAN METODE BACKPROPAGATION," 2016.
- [11] Y. D. Lestari, "Jaringan syaraf tiruan untuk prediksi penjualan jamur menggunakan algoritma backpropagation," *J. ISD*, vol. 2, no. 1, pp. 40–46, 2017.
- [12] I. Menarianti, "Klasifikasi Data Mining Dalam Menentukan Pemberian Kredit Bagi Nasabah Koperasi," *Ilm. Teknosains*, vol. 1, no. 1, pp. 36–45, 2015.

Lampiran

Kode		Komentar	Status
0	1	semoga kedepannya bisa lebih baik lagi	1
1	2	semoga lebih bisa diperbaiki lagi yang masih k...	1
2	3	tingkatkan	1
4	5	bagus	1
5	6	sudah cukup baik tingkatkan lagi	1
6	7	lebih diperbaiki lagi	1
7	8	cukup baik lebih ditingkatkan lagi	1
8	9	belum pernah ngurusin ke lak	0
9	10	harap di tingkatkan	1
10	11	jadilah lebih baik	1
11	12	sudah cukup	1
12	13	kurang cepat menanggapi	0
13	14	bagus semoga layanan akademik nya lebih baik d...	1
14	15	sudah baik	1
15	16	semoga lebih baik kedepannya	1
16	17	agar menjadi lebih baik kedepannya	1
17	18	tetap kepentingan permasalahan mahasiswa denga...	0
18	19	baik	1
19	20	bapak ibu maaf sebelumnya saya adalah alumni s...	1
20	21	semoga pihak layanan akademik dapat mempertahankan...	1
21	22	kurang efektif waktu matakuliah pilihan jam la...	0
22	23	lebih di tingkatkan lagi proyektor banyak yang...	1
23	24	oke	1
24	25	lbh baik lg kedepan	1
25	26	sangat baik sejauh ini hanya saja tidak sediki...	0
26	27	semoga dipermudah	1
27	28	sudah baik namun perlu ditingkatkan lagi	1
28	29	yang harus di lakukan adalah tetap menjaga kua...	1

Lampiran 1. Hasil Prediksi